# Testing the stability of an acquiescence style factor behind two interrelated substantive variables in a panel design

Billiet, Jaak B ; Davidov, Eldad

Abstract: This article addresses the question of to what extent one type of response style, called acquiescence (or agreeing response bias), is stable over time. A structural equation modeling approach is applied to measure the stability of one acquiescence factor behind two concepts among the same respondents for a 4-year period. The data used are representative population surveys in 1995 and 1999 from the Belgian Election Study in which balanced sets of items are used for measuring two interrelated constructs: perceived ethnic threat and distrust in politics. This study provides empirical support that acquiescence is stable and consistent for a 4-year period.

DOI: https://doi.org/10.1177/0049124107313901

# TESTING THE STABILITY OF AN ACQUIESCENCE STYLE FACTOR BEHIND TWO INTERRELATED SUBSTANTIVE VARIABLES IN A PANEL DESIGN

Jacques Billiet, Eldad Davidov

Centre for Sociological Research, Katholieke Universiteit Leuven

*This is a pre-copy-editing, author-produced PDF following peer review. The final, definitive version of this paper has been published in* **Sociological Methods & Research** Vol 36(4), May 2008, 542–562 *by SAGE Publications Ltd (All rights reserved. ©) and is available online under* **doi:10.1177/0049124107313901**

# TESTING THE STABILITY OF AN ACQUIESCENCE STYLE FACTOR BEHIND TWO INTERRELATED SUBSTANTIVE VARIABLES IN A PANEL DESIGN

Jaak B. Billiet*

Eldad Davidov

*Contact address: Centre for Sociological Research, Katholieke Universiteit Leuven, E. van Evenstraat 2B, 3000 Leuven, Belgium. Email: Jaak.Billiet@soc.kuleuven.be

# TESTING THE STABILITY OF AN ACQUIESCENCE STYLE FACTOR BEHIND TWO INTERRELATED SUBSTANTIVE VARIABLES IN A PANEL DESIGN

Jaak B. Billiet

Eldad Davidov

Centre for Sociological Research

Katholieke Universiteit Leuven

*Abstract.* This article addresses the question as to what extent one type of response style called acquiescence (or agreeing response bias) is stable over time. A structural equation modelling (SEM) approach is applied to measure the stability of one acquiescence factor behind two concepts among the same respondents over a four-year period of time. The data used are representative population surveys in 1995 and 1999 from the Belgian Election Study in which balanced sets of items are used for measuring two interrelated constructs: perceived ethnic threat and distrust in politics. This study provides empirical support that acquiescence is stable and consistent over a four-year period of time.

Key Words: acquiescence, stability, structural equation modelling, Belgian Election Study, perceived ethnic threat, distrust in politics

Surveys often use multiple indicators in order to measure latent variables such as attitudes, opinions, or beliefs. The response scales utilized to measure these indicators are often written in a "Likert" format. In this format, interviewees are requested to answer to what extent they agree or disagree with different statements. Researchers have promoted the use of multiple indicators for these types of questions to account for random measurement error. However, such a response format may be susceptible to systematic errors as well (Saris, 1998). Whereas one part of the systematic error of measurement is content related, and may be a result of social desirability or other determinants creating a false image related to the content of the question items, some studies have shown that this response format is also susceptible to a non-content related component of the systematic error, the so-called response style effect (McClendon 1991a,b; Weijters 2006). One form of style effects is an agreeing-response bias or acquiescence. Acquiescence is the tendency to answer items in a positive way regardless of their content. In other words, it is the tendency of some respondents to agree with an item irrespective of the content of that item (Bentler, Jackson, and Messick 1971; Cloud and Vaughan 1970; Hamilton 1968; Ray 1984a,b; Watson 1992).

Not accounting for acquiescence may distort the item-construct relations and the relations between content variables and predictors in a model (Cheung and Rensvold 2000; Baumgartner and Steenkamp 2001). In comparative research, acquiescence may lead to biases in the assessment of the invariance of loadings of content factors across groups or countries (Cheung and Rensvold 2000; Welkenhuysen-Gybels, Billiet, and Cambré 2003:708-710). The same applies to comparative research with repeated cross-sectional samples over time. In models based on panel data such as autoregressive models, acquiescence may lead to distortions in the estimation of stability coefficients and cross-lagged effects between variables over time. In the research field of method effects, this topic has received considerable attention (e.g., Bentler, Jackson, and Messick 1971; Cloud and Vaughan 1970;

Hamilton 1968; McClendon 1991a,b; Ray 1984a,b; Toner 1987). Most of these studies were concerned with measuring and controlling for acquiescence in their data.

Two sources for acquiescence are discussed in previous studies (Weijters 2006:29-31): the stimulus (e.g., the question posed to the individual) and the respondent. One research tradition has focussed on the first source, and especially on how to optimise the question wording so as to reduce acquiescence and other style effects (e.g., Schuman and Presser 1981). The second tradition has focussed on the respondent's personality as a major determinant of style effects. The personality is considered to have a certain orientation towards question wording, which influences the level of acquiescence (e.g., Bentler, Jackson, and Messick 1971). This second view has been criticised by several authors. For example, Rorer (1965) rejects the view that acquiescence is an important personality trait because measures of it fail to correlate across tests. Rorer hypothesised that response styles are test specific. As such, he also does not expect them to have any stability of acquiescence over time. However, according to the second tradition, the question of whether acquiescence is stable or not over time must be answered in an empirical way. Further controversies on the importance and identification of acquiescence have been delineated, for example, in Bentler, Jackson, and Messick (1971) as well as in Block (1971). Bentler et al. (1971:187-190) offered empirical evidence from psychological tests that acquiescence was best understood as being composed of at least two different processes, one related to social aphorisms (agreement acquiescence) and one related to personality inventories (acceptance acquiescence).

In this study we use an SEM approach to measure the stability of agreement acquiescence among the same respondents over a four-year period of time in the population surveys of the 1995 and 1999 Belgian Election Studies. Measurement models for two balanced sets of items measuring two related constructs, perceived ethnic threat and distrust in politics, are tested to answer the following question: Is agreement acquiescence consistent over time when the same measurements of the same constructs are used among the same

respondents? In contrast to Rorer's (1965) postulation that response styles are trivially important, this study provides empirical support that acquiescence is highly stable and consistent across a four-year period of time.

## PREVIOUS WORK AND RESEARCH QUESTION

Different methods for measuring or controlling acquiescence have been advocated. Martin (1964), Cloud and Vaughn (1970), and Ray (1979) proposed the use of balanced scales as a simple and effective method for controlling acquiescence. In this method, half of the items are worded in the positive direction and half in the negative direction. The idea justifying balanced items is that acquiescence to the positively worded items will be cancelled out by acquiescence to the negative items (McClendon 1992). As a result, the scale mean will not be biased due to acquiescence (but may be biased due to other method effects). Furthermore, its variance will not be affected by acquiescence as correlations between items worded in the same direction will increase, (which will increase the variance of the scale), and correlation between positively worded items and negatively worded items will decrease, (which will decrease the variance of the scale) (McClendon 1992). As a result, the co-variance between the balanced scale and another variable will also not be biased to both negative and positive items. However, the assumption underlying the justification for balanced scales is that positively and negatively worded items are equally susceptible to acquiescence, but this is not always the case, as McClendon (1992) demonstrated. Another characteristic of acquiescence-balanced scales is that their alpha reliability coefficients often seem low. On the other hand, when scales are not balanced, acquiescence tendencies artificially inflate the alpha reliability coefficients. In contrast, to avoid this assumption, Schuman and Presser (1981:207) and Toner (1987) suggested using forced-choice questions in place of agree-disagree items. This

suggestion is not often followed by researchers who want to use latent variables based on multiple indicators. There are other reasons for researchers rejecting forced-choice questions. First, this type of questions takes two statements to produce a dichotomous indicator with an unusually low yield of information per statement. Secondly, respondents find forced choices annoying if they see the statements as not pure opposites. Finally, forced-choice questions sound clumsy if read to respondents as, for example, in telephone interviews.

In an important study, Mirowsky and Ross (1991) proposed the use of structural equation modelling (SEM) to control for acquiescence. They specified a model that allowed for the measurement and control of acquiescence with a balanced set of items. The items were used to measure sense of control. In this model, each of the control items was specified as having a fixed equal (positive) loading on the factor acquiescence. These items also loaded on the sense of control factor. Using population samples, they found a significant variance for the acquiescence factor and interpreted it as the acquiescence underlying these items. Furthermore, they used several criteria to validate their measurement model, such as the correlation of acquiescence with other variables (age or education) and the goodness-of-fit of the model. However, it should be noted that acquiescence was not directly tested or measured, but rather only modelled in their study.

Several other approaches for the detection of acquiescence have been suggested. Cheung and Rensvold (2000) propose using SEM to detect acquiescence across groups. In their approach, a model in which the item intercepts ($\tau_i$) are set equal across (cultural) groups is tested against a model without these equality constraints. A significant difference in the fit of these two models indicates significant differences between the item intercepts, which leads to the conclusion that at least some of the items are influenced by an acquiescent response bias. A major drawback of this procedure is the fact that item intercept differences across groups can just as well indicate real attitudinal differences between the groups as differences due to the Likert format of the items.

In the field of marketing research, some scholars (Baumgartner and Steenkamp 2001; Weijters 2006), following a method proposed by Greenleaf (1992a,b), make use of a set of items that is maximally heterogeneous in content. The basic idea behind this approach is to reduce the effect of content in the set of items to random noise: If all the items represent different constructs that are (on average) unrelated, it can be expected that there is no consistency in responses other than that induced by response styles (Weijters 2006:42). Acquiescence is then measured by the sum of agreements. Other response styles are also measured by this method: extreme response style, the tendency to choose the middle of the scale, and even the tendency to reject the items. The advantage of Mirowsky and Ross' (1991) SEM approach over that proposed by Greenleaf is that no large extra set of unrelated items is needed in order to account for a style effect. The disadvantage of the balanced set of items approach is that it only applies to acquiescence (or to dis-acquiescence) and not to other response styles.

Following the study of Mirowsky and Ross (1991), Billiet and McClendon (2000) measured acquiescence for two balanced sets of items measuring perceived ethnic threat and distrust in politics. The estimation of a model with two content factors (for the two sets of items measuring perceived ethnic threat and distrust in politics) and a style factor (acquiescence) which loaded equally on all items yielded a significantly better fit than a model with only two content factors (see Figure 1). The two content factors correlated only minimally with the style factor thus confirming their theoretical postulation, which did not expect any substantial correlations between content factors and a response style or a method effect. The study provided strong support for the existence of acquiescence, particularly because it was conducted on two independent samples of the Flemish population in Belgium as well as on a representative sample from Wallonia in 1995.

Figure 1 about here

To identify the style factor and investigate what it stands for and whether it actually represents acquiescence, Billiet and McClendon (2000) explored its correlations with education and age. Less educated respondents may have less clear views on certain types of items and, therefore, have a higher tendency to display acquiescence (McClendon 1991b; Mirowsky and Ross 1991; Schuman and Presser 1981). Moreover, Mirowsky and Ross (1991) found that acquiescence increases with age. As theoretically expected, the results of the external validation yielded a negative association with education and a positive association with age. Using the same method for detecting acquiescence in a multi-group comparison between respondents with high and respondents with low political knowledge, Billiet, Swyngedouw, and Waege (2004) found that the latter were more subjected to acquiescence in a political alienation scale than the better informed respondents.

To obtain a much stronger validation, Billiet and McClendon (2000:623-626) investigated the relation between the style factor and a variable that simply measures the sum of agreements across a balanced set of 14 positively and negatively worded items. The latter is called "scoring for acquiescence" (Ray 1979). The 14 items consist of the 10 indicators that were used for measuring perceived ethnic threat and distrust in politics plus four additional indicators measuring utilitarian individualism and communalism in the 1995 survey. This procedure led to a strong correlation of 0.90 ($t = 22.26$) between "scoring for acquiescence" and the style factor in a random sample of Flemish voters in Belgium. This finding has been replicated in an independent random sample of the Walloon population in the 1995 General Elections Survey in Belgium and in other samples since then (Billiet, Cambré, and Welkenhuysen-Gybels 2002). It was concluded that both the latent style factor and "scoring for acquiescence" were measuring the same thing, namely, a tendency to endorse statements irrespective of their content.

Some studies have focussed on the measurement of acquiescence in cross-cultural surveys (Billiet, Maddens, and Beerten, 2003; Cambré, Welkenhuysen-Gybels, and Billiet 2002; Billiet et al. 2002; Welkenhuysen-Gybels et al. 2003). In these cases, modelling a style factor led more easily to the acceptance of adequate cross-country equivalent measurement models. It was found that the respondents in some countries were somewhat more affected by acquiescence than in other countries. Contrary to expectations, none of the relations between the content variables changed after inclusion of the style factor. This is in reality, however, not necessarily the case. Moors (2004), for example, used a balanced set of items to measure perceived discrimination in several ethnic minority groups but, contrary to our confirmatory factor analysis approach, he used latent class models to detect a latent class of acquiescent respondents. He concluded that no 'true' difference between ethnic groups on 'perceived discrimination' is found after controlling for response style (Moors, 2004:317).

The measurement of response style is useful not only in cross-cultural surveys but also in studies of change over time. In this respect a serious question still remains: Is acquiescence stable among respondents over time? This question is relevant for the purpose of concluding whether the 'style' factor that was measured might be conceived as a response style effect which is more dependent on respondent (personality) traits than on question characteristics. In this study we are going to test the proposition that, among the same respondents, acquiescence is stable over time.

Critical views, which do not consider acquiescence as a personality trait, do not expect it to be stable over time nor to be common behind measurements of two or more content factors. For instance, Rorer (1965) argues that there is no evidence suggesting that response style is a personality trait of any importance. He concludes in his study that response styles are only trivially important in determining responses to personality or attitude questions in a survey. Ray (1983) argues that acquiescence is influenced by the ambiguity of certain measurements and not by individuals. Therefore, he expected to have different levels of acquiescence for

different items. Furthermore, he argued that individuals who may display acquiescence to certain items may not display it when responding to other items. To establish the existence of a stable response style, one has to show that it applies to different sets of items and remains stable over time.

In spite of such contra-arguments, several scholars have successfully demonstrated the existence of acquiescence in different content domains (see, e.g., Baumgartner and Steenkamp 2001; Paulhaus 1991). However, so far there have been almost no empirical investigations of the stability of acquiescence over time. Recently, Weijters (2006) assessed the extent to which individual response styles are stable for the same individuals across a period of time of one year. He found a substantial level of stability for the style factor. To the best of our knowledge, no other assessments of stability were conducted over longer periods of time or with representative population data. In this study we are going to test the proposition that the acquiescence of panel respondents measured in 1995 is stable over time using data from the 1995 and 1999 waves of the Belgian Election Study. We will try to identify the acquiescence factors in each wave, and evaluate the co-variance (or correlation) between the two style factors over a four-year period (1995-1999). A substantially high correlation indicates a high level of stability, but a correlation that is not significantly different from zero will indicate that one cannot predict acquiescence in 1999 on the basis of acquiescence information from 1995 at the individual level.

## DATA

In this study we utilize the responses to two balanced sets of items that were collected in 1995 and in 1999 in the Belgian Election Study. Respondents in this study participated in a three wave panel survey of the general elections that took place in Belgium in 1991, 1995, and 1999. In order to correct the samples of 1995 and 1999 for panel attrition and to make them

random samples of the population (i.e., of the Flemish voters at each election) the panel parts of 1995 and 1999 were re–freshed by randomly selected respondents at each occasion. We are not interested in the original sample from1991 since these data did not contain the balanced scales that are used in this study. Only the samples of Flemish speaking voters are used because the samples of the French speaking voters do not contain usable panel data (for further details on the sample, see Billiet, Swyngedouw, Carton, and Beerten, 1997, and Billiet, Swyngedouw, Depickere, and Meerseman, 2000).[1] The response rate in 1999 of the panel respondents who take part in both the 1995 and the 1999 surveys (N = 1,503) was 68.8 percent of the size of the sample in the 1995 survey (N = 2,184).

It is very unlikely that our findings are biased in the desired direction of what we want to show because of panel attrition. First, the parameters of the style factors that were found in the panel part of the dataset are similar to the style factors parameters that were previously detected in the complete cross sections of the 1995 and 1999 surveys (see Billiet, Maddens, and Beerten, 2003). Furthermore, the joint distribution according to gender, age categories (six classes), and education (three levels) of the panel respondents who participate in both waves do not significantly differ from the total 1995 sample (Chi-square = 30.421; df = 25; p = 0.209). The deviation between the joint distribution (of gender, age, and education) in the panel part and the joint distribution in the total sample of 1999 (N = 2,305) is larger (Chi-square = 132.335; df = 25; $p < 0.001$). The reason for this is that fresh voters were randomly included in the complete sample in order to optimally represent the population of voters in 1999, and to correct for panel attrition. As a consequence of the panel attrition, senior voters (aged 65 years and older) and lower educated voters are underrepresented in the panel part compared to the complete 1999 sample. It is, however, not very likely that this artificially inflates our findings because lower educated and older respondents are more subjected to acquiescence than others. This may rather lead to an underestimation of the size of variance of the style factor.

The questionnaire contained two balanced sets of items, measuring distrust in politics and perceived ethnic threat. The 5-point response scales ranged from 1 (*strongly agree*) to 5 (*strongly disagree*) in the questionnaires, but the scores were all reversed in the analysis so that strongly agree obtained the highest score (5) and strongly disagree the lowest (1). About 26 percent of the respondents in 1995 and 24 percent of the respondents in 1999 agreed with at least one pair of quasi-contradictory items. Table 1 provides a description of the 10 items that we used. For determining perceived ethnic threat, three negative items measure the extent to which immigrants, according to the respondents, are not to be trusted, endanger the employment of the Belgians, and are a threat to Belgian culture and customs. Three positive items measure the extent to which, from the respondents' point of view, immigrants contribute to the prosperity of Belgium, enrich Belgian society, and are wholeheartedly welcome. These three pairs of positively and negatively worded items are not pure reversals, and it is logically possible to agree to some extent with both of the items in each pair, but it is not very likely that an individual would endorse two or all three pairs if one has a consistent attitude toward immigrants. For measuring distrust in politics, two negative items measure the extent to which, according to the respondents, politicians have lost the ability to listen to ordinary people and the extent to which politicians feel themselves too good for ordinary people once they are elected. Two positive items measure the extent to which respondents agree that their views are generally taken into account and that politicians are capable people who know what they are doing.

Table 1 about here

## METHOD

The origin of model testing is a model of the 1995 and 1999 data in which only two content factors – and no style factor – are specified. In the next step, this model is compared with the

second model in which an additional style factor is specified.[2] The models are estimated with LISREL8.54® (Jöreskog and Sörbom 1996) using an asymptotic co-variance matrix of polychoric correlations and the weighted least squares (WLS) estimation procedure. The WLS procedure uses an asymptotically distribution free function and was proposed by Jöreskog (1990) for ordinal data and when sufficiently large samples are available (Bollen 1989:426; Jöreskog and Sörbom 1993). A consistent estimation of the asymptotic co-variance matrix of polychoric correlations was computed using PRELIS2 (Jöreskog and Sörbom 1993:25-28). Because of the preference for a WLS estimation procedure, which requires list-wise deletion of cases with missing values in the 20 items (10 in 1995 and 10 in 1999), the actual sample size is reduced to 1,112 units. Analysis of item non-response shows that it is very unlikely that lost cases have an effect on our main conclusions.[3]

In order to obtain measurement invariance over time, the corresponding regression coefficients between items in 1995 and 1999 and their content factors are constrained to be equal over time. In this way, the procedure proposed by Rensvold and Cheung (1998) and by Welkenhuysen-Gybels et al. (2003) to control for factorial invariance in measurement models across cultural groups is applied to measurement models in a single cultural group observed on several occasions (for an application, see Schlüter, Davidov, and Schmidt 2006). Furthermore, only residual co-variances between corresponding (identical) items in 1995 and 1999 are allowed to correlate. The model specifications are shown in the Appendix.

## RESULTS

Table 2 presents the three models that are tested. In the first model we include only two content factors for each occasion of measurement. The co-variance between the content variables at one time point and over time is estimated since, on theoretical grounds, it is assumed that these co-variances are different from zero (see Appendix). This basic model is

tested against two models with style factors. In the second model two style factors are assumed, one for the 1995 data and one for the 1999 data. Each factor loads equally (denoted by the number 1 in Figure 2) on the items of the respective year because it is assumed that in each year all the items of the two content factors are equally subject to the same style effect. This style effect was confirmed in a previous study (Billiet and McClendon 2000) for the 1995 panel wave, and it is now introduced for the 1999 wave. In Model 2, the style factors are not allowed to correlate (see specifications in Appendix). In other words, we assume that there is no stability at all in the style factor. In the third model, the style factors are allowed to correlate, but the specification that no correlation is allowed between style and the two content factors remains (see Appendix). In this way the stability of the style (acquiescence) factor across time is tested. Model 3 is displayed in Figure 2 and the results are reported in Table 2.

Figure 2 about here

To assess model fit, several criteria are used (for an overview see Hu and Bentler 1998). The models are evaluated by examining the value of the root mean square error of approximation (RMSEA) and the *p* value of close fit (Pclose). These two parameters of global fit have been recommended by several researchers to discern between well-fitting and poorly fitting models (Shevlin and Miles 1998; Billiet and McClendon 2000). A model that fits the data well will have a RMSEA value smaller than 0.05 and a Pclose value larger than 0.5 (Browne and Cudeck 1993). In our case, however, since the sample size is large and the *p* value may then reject models with small misspecifications (Saris, Satorra, and Sörbom 1987; Saris and Satorra 1993), we cannot rely on the *p* value in order to select a model. In our analyses, we will see that none of the three models can be rejected on the basis of the criteria that have been mentioned here. Therefore, we also examine the modification index of the

parameter under study and the decrease in the chi-square value for the loss of one degree of freedom. This latter criterion is suitable since Model 3 is nested in Model 2, which in turn is nested in Model 1.[4]

In Model 1, the RMSEA and Pclose values indicate a good fit (RMSEA = 0.037 and Pclose = 1.0). However, on the basis of the difference in chi-square, Model 2 is preferable to Model 1. The introduction of a style factor for 1995 and 1999 in Model 2 produces a significant drop in chi-square of 48.37 for a loss of 2 degrees of freedom. In Model 2, the Pclose and the RMSEA values still indicate a good fit with a slight drop in RMSEA (RMSEA = 0.033, Pclose = 1.0). Model 3 allows a correlation between the style factor in 1995 and the style factor in 1999. This change improved the model fit significantly and caused a further drop in chi-square of 13.04 for a loss of one degree of freedom (RMSEA = 0.032 and Pclose = 1.0). According to the fit indices, Model 3 is the preferred model.

Table 3 summarises the standardized regression coefficients between the content and style factors and the indicators in Model 3. Perceived ethnic threat in 1995 and in 1999 (THR95 and THR99) load strongly on their corresponding items. The loadings are positive for the negatively worded items and negative for the positively worded items. Distrust in politics in 1995 and in 1999 (DISTR95 and DISTR99) also load quite strongly on their respective indicators, with positive loadings on the negatively worded items and negative loadings on the positively worded items. The two style factors in 1995 and in 1999 (STYLE95 and STYLE99) load equally on the items.

Table 3 about here

Table 4 reports the correlations between the content factors, between the style factors, and between the content and style factors. The correlation between the style factors was high ($r = 0.562$, t-value = 3.612, $p < 0.05$) and – taking into consideration the long four-year period

between the panel waves – it indicates a high stability in response style bias among our panel respondents. This correlation, however, was lower than between the content factors over time, which displayed an even higher stability ($r_{(THR95,THR99)} = 0.859$; $r_{(DISTR95,DISTR99)} = 0.736$). The correlation between the content factors perceived ethnic threat and distrust in politics in 1995 ($r = 0.496$) and in 1999 ($r = 0.592$) was not influenced by the introduction of the style factors ($r_{(model\ 1,95)} = 0.507$, $r_{(model\ 1,\ 99)} = 0.596$), and is strong and positive as expected according to theory (Billiet 1995). The zero correlations between the style and content factors are in line with our expectations. There is no theoretical reason for perceived ethnic threat and political efficacy to correlate with style. However, it does not necessarily imply a null correlation in other circumstances. It may be the case that the counterbalancing of items suppresses the correlation.

Table 4 about here

## DISCUSSION AND CONCLUSIONS

A common style factor underlying two balanced sets of Likert-type items is identified using a representative data set of the Flemish population in Belgium. Previous studies have shown that this common method factor correlates very strongly with the number of times that the respondents agree with the items in the balanced sets. For this reason this response style is called ‘acquiescence’. Furthermore, we have shown how acquiescence can be identified and controlled in panel data in which identical balanced sets of items are measured on two occasions over a long period of time.

The selection of Model 3 was based on theoretical as well as empirical grounds. Model 3, which includes a method factor for acquiescence, fits the data much better than Model 1, which does not include a method factor. Furthermore, Model 3, which specifies a non-zero

correlation between the style factors in 1995 and 1999, fits the data better than Model 2, in which this correlation is fixed to zero. In fact, the stability of acquiescence was not larger than the stability of distrust in politics and perceived ethnic threat, but was still considerable when we take into account the long period of time between the two measurement occasions. In other words, the findings of this study strongly support the hypothesis that acquiescence is stable over time.

In future research it would be useful to identify acquiescence and to test its stability over more than two occasions of measurement. This would enable us to estimate the possible development and growth in acquiescence over time using latent growth modelling. Furthermore, one could test to what extent the initial level of acquiescence and its trajectory of change may be related to content variables.

The fact that a previous study (Billiet and McClendon 2000) found a high correlation between the style factor and the number of positive answers to different question items supports, to a large extent, the assumption that what we identified is an acquiescence factor. Our results do not guarantee, however, that the style factor identified truly represents acquiescence, in spite of the validation tests that were conducted. Rather than a yes-saying tendency, it could also reflect a tendency to use the last category on the scale or a preference for an extreme response category. Providing a better answer for this issue should involve a new design of the study with various orders of response categories or the inclusion of response categories that do not refer to agreement or disagreement in the balanced set of five-point Likert items.[5]

We expected to find substantial changes in the results after introducing the acquiescence factor over time, arguing that models with such a method factor produce more valid results (Billiet and McClendon 2000). In our case, controlling for acquiescence did not change the correlations between the content factors, but this was not our goal. It may be the case here that as the content domains may be emotionally hot ones, no change in the correlations between

the content factors was observed. It is also possible that acquiescence to the positively worded items was cancelled out by acquiescence to the negatively worded items. The strategy of balancing the items contextually suppressed a bias that otherwise would have altered the correlation. Yet the acquiescence response exists even in this context.

The question that arises is whether it is worthwhile to model acquiescence in research practice. We could defend this modelling strategy by arguing that substantive results may be seriously biased in other situations with a different set of content variables if a style factor is not included, as has been shown by McClendon (1992). But even when it does not change the substantial results, it is important to include it because it represents an important part of the model. Survey questions are subjected to systematic errors and it is mostly not possible to reduce or avoid these errors during the data collection. Modelling them instead is a practical solution. Often, models that are thought to include several substantive latent variables eventually retain only one substantive latent variable when a style factor is introduced (see, e.g., Marsh 1996:810; van Schuur and Kiers 1994). Moreover, in cross-cultural data, omitting a factor which accounts for acquiescence may bias the assessment of measurement invariance of content factors across the cultural groups under study (see Welkenhuysen-Gybels et al. 2003). Its impact may actually differ between countries. The different studies illustrate that introducing an acquiescence factor to control for response bias in balanced Likert-type items not only improves the model fit in both cross-sectional, longitudinal and cross-national studies, but also increases parsimony and simplifies interpretation. This is especially important in panel designs as well as in cross-cultural research, where models tend to become complicated because of large data sets and a higher number of constructs for different time points or groups. As was demonstrated here, the factor of acquiescence seems to be a real trait that varies across persons, it is relatively stable over time, and accounting for it provides researchers with a useful way to deal with this type of style effect in survey research.

**Acknowledgements**

The second author would like to thank the Katholieke Universiteit Leuven for the Postdoctoral Fellowship that supported this study. The two authors would like to thank the editor and two anonymous reviewers for their constructive comments. Many thanks to Lisa Trierweiler for the English-proof of the manuscript.

REFERENCES

Baumgartner, Hans and Jan-Benedict E.M. Steenkamp. 2001. "Response Styles in Marketing Research: A Cross National Investigation." *Journal of Marketing Research* 38: 143-156.

Bentler, Peter M., Douglas N. Jackson, and Samuel Messick. 1971. "Identification of Content and Style: A Two Dimensional Interpretation of Acquiescence." *Psychological Bulletin* 76: 186-204.

Billiet, Jaak. 1995. *Theoretical Dimensions and Measurement Models of Attitudes Ethnic Minorities.* Paper presented at the 8th conference of the International Research Group on Methodology and Comparative Survey Research, Prague.

Billiet, Jaak, Bart Cambré, and Jerry Welkenhuysen-Gybels. 2002. "Equivalence of measurement instruments for balanced sets of items in cross-cultural surveys, taking method effects into account." Pp. 53-72 in *Developments in Social Science Methodology*, edited by A. Ferligoj and A. Mrvar. Ljubljana: FVD.

Billiet, Jaak, Bart Maddens, and Roeland Beerten. 2003. "National Identity and Attitude Toward Foreigners in a Multinational State: A Replication." *Political Psychology* 24(2): 241-257.

Billiet, Jaak B. and McKee J. McClendon. 2000. "Modeling Acquiescence in Measurement Models for Two Balanced Sets of Items." *Structural Equation Modeling* 7(4): 608-628.

Billiet, Jaak, Marc Swyngedouw, Ann Carton, and Roeland Beerten. 1997. *1995 General Election Study Belgium - Flanders. Codebook and Questionnaire.* Leuven: ISPO.

Billiet, Jaak, Marc Swyngedouw, Astrid Depickere, and Erik Meerseman. 2000. *1999 General Election Study Belgium - Flanders. Codebook and Questionnaire.* Leuven: ISPO.

Billiet, Jaak, Marc Swyngedouw, and Hans Waege. 2004. "Attitude Strength and Response Stability of a Quasi-Balanced Political Alienation Scale in a Panel Study." Pp. 268-292, in *Studies in Public Opinion. Attitudes, Nonattitudes, Measurement Error, and Change*, edited by W.E. Saris and P.M. Sniderman. Princeton: Princeton University Press.

Block, Jack. 1971. "On Further Conjectures Regarding Acquiescence." *Psychological Bulletin* 76: 205-210.

Bollen, Kenneth A. 1989. *Structural Equation Models with Latent Variables.* New York: Wiley.

Browne, Michael. W. and Robert Cudeck. 1993. "Alternative Ways of Assessing Model Fit". Pp. 445-455 in *Testing Structural Equation Models*, edited by Kenneth. A. Bollen and J.Scott Long. Newsbury Park, CA: Sage.

Cambré, Bart, Jerry Welkenhuysen-Gybels, and Jaak Billiet. 2002. "Is It Content or Style? An Evaluation of Two Competitive Measurement Models Applied to a Balanced Set of Ethnocentrism Items." *International Journal of Comparative Sociology* 43(1): 1-20.

Cheung, Gordon W. and Rensvold, Roger B. 2000. "Assessing Extreme and Acquiescence Response Sets in Cross-Cultural Research Using Structural Equation Modelling." *Journal of Cross-Cultural Psychology* 31: 187-212.

Cloud, Jonathan and Graham M. Vaughan. 1970. "Using Balanced Scales to Control Acquiescence." *Sociometry* 33(2): 193-202.

Greenleaf, Eric A. 1992a. "Improving Rating Scale Parameters by Detecting and Correcting Bias Components in Some Response Styles." *Journal of Marketing Research* 29(2): 176-188.

Greenleaf, Eric A. 1992b. "Measuring Extreme Response Style." *Public Opinion Quarterly* 56(2): 328-350.

Hamilton, David L. 1968. “Personality Attributes Associated with Extreme Response Style.” *Psychological Bulletin* 69(3): 192-203.

Hu, Li-tze and Peter M. Bentler. 1998. “Fit Indices in Covariance Structure Modeling: Sensitivity to Underparameterized Model Specification”. *Psychological Methods*. 3: 424-453.

Jöreskog, Karl G. (1990). “New Developments in LISREL: Analysis of Ordinal Variables Using Polychoric Correlations and Weighted Least Squares.” *Quality and Quantity* 24: 387-404.

Jöreskog, Karl G. and Dag Sörbom (1993). *New Features in PRELIS2*. Chicago: Scientific Software International.

Jöreskog, Karl G. and Dag Sörbom (1996). *LISREL8.53 User’s Reference Guide.* Chicago: Scientific Software International.

Marsh, Herbert, W. (1996). “Positive and Negative Global Self-Esteem: A Substantively Meaningful Distinction or Artifactors?” *Journal of Personality and Social Psychology* 70: 810–819.

Martin, John. 1964. “Acquiescence: Measurement and Theory.” *British Journal of Social and Clinical Psychology* 3: 316-326.

McClendon, McKee J. 1991a. “Acquiescence and Recency Response-Order Effects in Interview Surveys.” *Sociological Methods and Research* 20: 60-103.

McClendon, McKee J. 1991b. “Acquiescence: Tests of the Cognitive Limitations and Question Ambiguity Hypotheses.” *Journal of Official Statistics* 7(2): 153-166.

McClendon, McKee J. 1992. *On the Measurement and Control of Acquiescence in Latent Variable Models.* Paper Presented at the Meeting of the American Sociological Association, Pittsburgh, PA.

Mirowsky, John and Catherine E. Ross. 1991. “Eliminating Defense and Agreement Bias from Measures of the Sense of Control: A 2x2 Index.” *Social Psychology Quarterly* 54(2): 127-145.

Moors, Guy. 2004. “Facts and Artefacts in the Comparison of Attitudes Among Ethnic Minorities. A Multigroup Latent Class Structure Model with Adjustment for Response Style Behavior”. *European Sociological Review*, 20(4): 303-320.

Paulhus, Delroy L. 1991. “Measurement and Control of Response Bias.” Pp. 17-59 in *Measures of Personality and Social Psychological Attitudes,* edited by J.P. Robinson, P.R. Shaver, and L.S. Wrightsman. San Diego: Academic Press.

Ray, John J. 1979. “Is the Acquiescent Response Style Problem Not So Mythical After All: Some Results from a Successful Balanced *F* Scale.” *Journal of Personality Assessment* 43: 638-643.

Ray, John J. 1983. “Reviving the Problem of Acquiescent Response Bias.” *Journal of Social Psychology* 121: 81-96.

Ray, John J. 1984a. “A Further Comment on the Winkler, Kanouse, and Ware Method of Controlling for Acquiescent Response Bias.” *Journal of Applied Psychology* 69(2): 359.

Ray, John J. 1984b. “Reinventing the Wheel: Winkler, Kanouse, and Ware on Acquiescent Response Set.” *Journal of Applied Psychology* 69(2): 353-355.

Rensvold, Roger B. and Gordon W. Cheung. 1998. “Testing Measurement Models for Factorial Invariance: A Systematic Approach.” *Educational and Psychological Measurement* 58: 1017-1034.

Rorer, Leonard G. 1965. “The Great Response-Style Myth.” *Psychological Bulletin* 63: 129-156.

Saris, Willem E. 1998. “The Efects of Measurement Error in Cross Cultural Research.” *ZUMA-Nachrichten Spezial* (January): 67-83.

Saris, Willem E. and Albert Satorra. 1993. "Power Evaluations in Structural Models." Pp. 181-204 in *Testing Structural Equation Models,* edited by Kenneth.A. Bollen and J. Scott Long. Newbury Park, CA: Sage.

Saris, Willem E., Albert Satorra, and Dag Sörbom. 1987. "The Detection and Correction of Specification Errors in Structural Equation Models." *Sociological Methodology* 17: 105-129.

Schlüter, Elmar, Eldad Davidov, and Peter Schmidt. 2006. "The Dynamics of Authoritarianism and Anomia: Applying Autoregressive Cross-Lagged and Latent Growth Models to a Three-Wave Panel Study." Pp. 315-336 in *Longitudinal Models in the Behavioral and Related Sciences,* edited by Kees van Montfort, Johan Oud, and Albert Satorra. EAM Book Series: Lawrence Erlbaum Publishers.

Schuman, Howard and Stanely Presser. 1981. *Questions and Answers in Attitude Surveys.* New York: Academic.

Shevlin, Mark and Jeremy Miles. 1998. "Effects of Sample Size Specification and Factor Loadings on the GFI in Confirmatory Factor Analysis." *Personality and Individual Differences* 21: 85-90.

Toner, Bill. 1987. "The Impact of Agreement Bias on the Ranking of Questionnaire Response." *Journal of Social Psychology* 127: 221-222.

Van de Vijver, Fons and Kwok Leung. 1997. *Methods and Data Analysis for Cross-Cultural Research.* Thousand Oaks: Sage.

Van Schuur, Wijbrandt. H. and Henk A. L. Kiers. 1994. "Why Factor Analysis Often Is the Incorrect Model for Analyzing Bipolar Concepts, and What Model to Use Instead." *Applied Psychological Measurement* 18: 97–110.

Watson, Dorothy. 1992. “Correcting for Acquiescent Response Bias in the Absence of a Balanced Scale: An Application to Class Consciousness.” *Sociological Methods and Research* 21: 52-88.

Weijters, Bert. 2006. *Response Styles in Consumer Research.* Belgium: Vlerick Leuven Gent Management School.

Welkenhuysen-Gybels, Jerry, Jaak Billiet, and Bart Cambré. 2003. “Adjustment for Acquiescence in the Assessment of the Construct Equivalence of Likert-Type Score Items.” *Journal of Cross-Cultural Psychology* 34(6): 702-722.

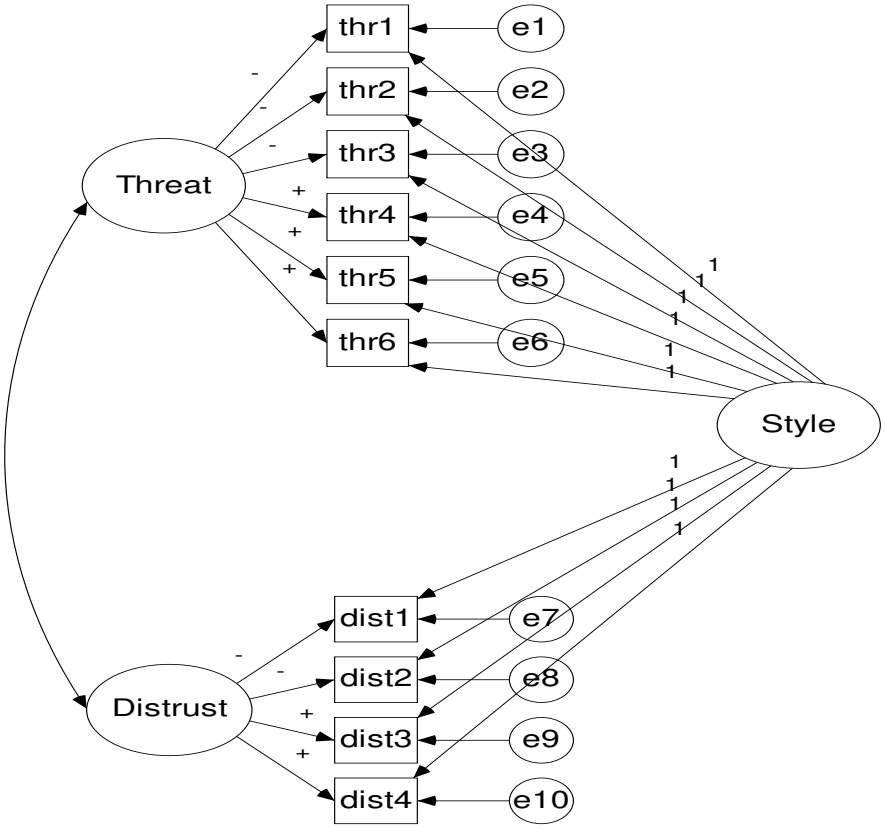



Figure 1: Measurement model for two interrelated content factors and a style factor. The model contains two content factors, Threat and Distrust, with balanced sets of six and four observed indicators respectively, and with identical slopes for the style factors (denoted 1).

(for actual parameter values, see Billiet and McClendon 2000:622-625).

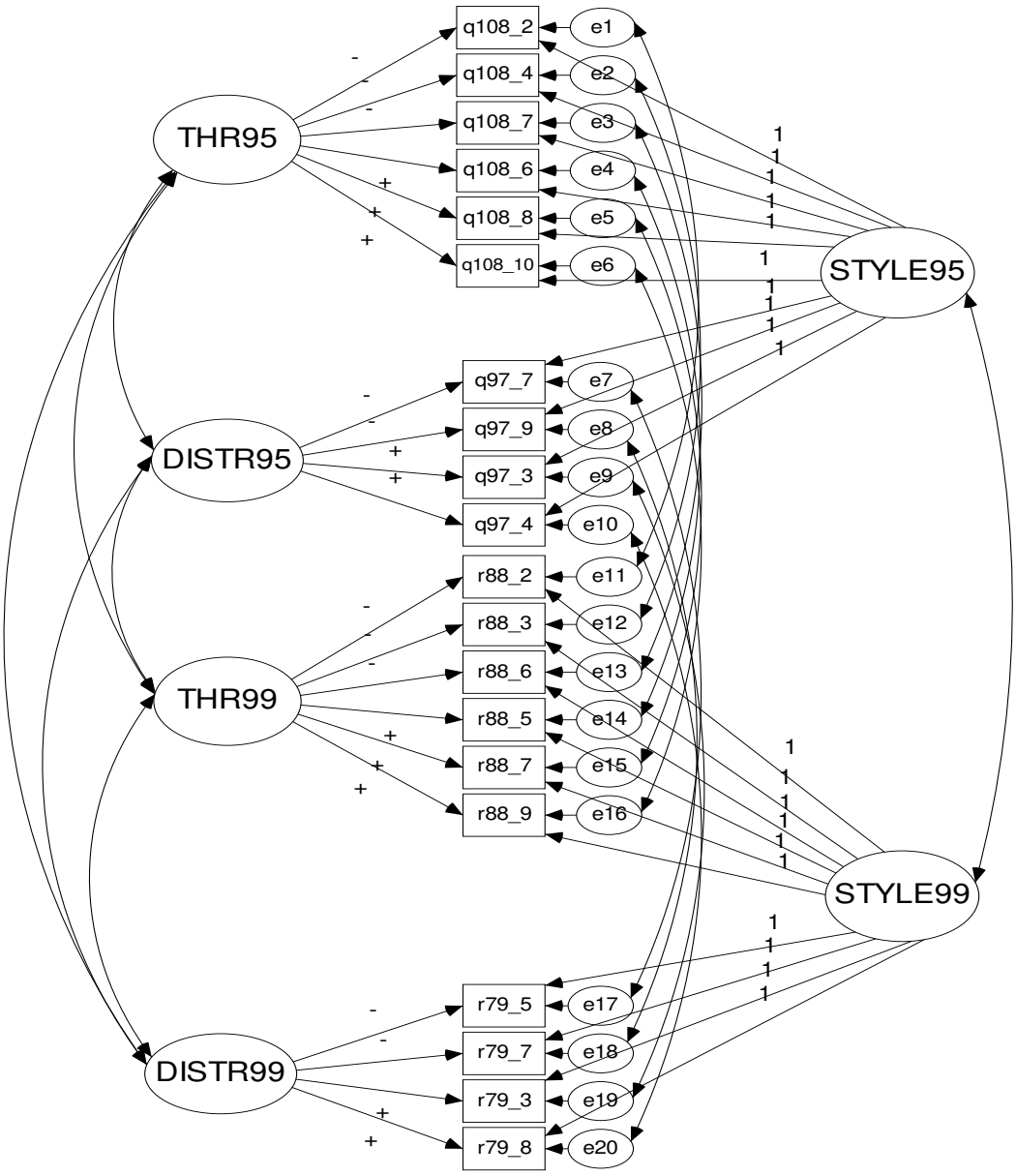



Figure 2: Measurement model for two content factors and one style factor in 1995 and in 1999. The model contains two content factors, Threat and Distrust, with balanced sets of six and four observed indicators, respectively, and with identical slopes for the style factors (denoted 1). The actual parameter values are reported in Tables 3 and 4.

Table 1: Balanced Sets of Items to Measure Perceived Ethnic Threat and Distrust in Politics in the Questionnaire of the 1995 and 1999 General Elections Survey

| 1995 | 1999 | Balanced sets of items |
|---|---|---|
| Threat | | |
| q108_2 | r88_2 | - In general, immigrants are not to be trusted. |
| q108_4 | r88_3 | - Guest workers endanger the employment of the Belgians. |
| q108_7 | r88_6 | - Muslims are a threat to our culture and customs. |
| q108_6 | r88_5 | + The immigrants contribute to the prosperity of our country. |
| q108_8 | r88_7 | + The presence of different cultures enriches our society. |
| q108_10 | r88_9 | + We should wholeheartedly welcome the foreigners who come to live here. |
| Distrust | | |
| q97_7 | r79_5 | - The politicians have lost the ability to listen to ordinary people like me. |
| q97_9 | r79_7 | - Once they are elected, most politicians feel themselves too good for people like me. |
| q97_3 | r79_3 | + If people like me make their views known, politicians generally take them into account. |
| q97_4 | r79_8 | + Most of our politicians are capable people who know what they are doing. |

Note. - negatively worded items; + positively worded items

Table 2: Models and their Indices of Model Fit (N = 1,112)

| Model | Chi-Square | DF | RMSEA | Pclose |
|---|---|---|---|---|
| Model 1: 2x2 content factors no style | 405.57 | 162 | 0.037 | 1.00 |
| Model 2: Style factors, cor(St95, St99)=0 | 357.20 | 160 | 0.033 | 1.00 |
| Model 3: Correlated style factors St95-St99 | 344.16 | 159 | 0.032 | 1.00 |

Note. t-value of cov (St95, St99) in Model 2 = 3.612 (P<0.05)

Table 3: Standardized Regression Coefficients (Factor Loadings) of Model 3 (T-Values in Parentheses).

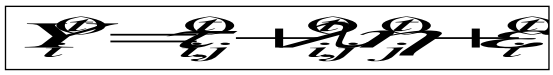
| | THR95 | DISTR95 | STYLE95 | THR99 | DISTR99 | STYLE99 |
|---|---|---|---|---|---|---|
| q108_2 | 0.835 (fixed) | --- | 0.180 | --- | --- | --- |
| q108_4 | 0.785 (55.36) | --- | 0.180 | --- | --- | --- |
| q108_7 | 0.703 (46.60) | --- | 0.180 | --- | --- | --- |
| q108_6 | -0.772 (-48.78) | --- | 0.180 | --- | --- | --- |
| q108_8 | -0.822 (-52.94) | --- | 0.180 | --- | --- | --- |
| q108_10 | -0.713 (-41.88) | --- | 0.180 | --- | --- | --- |
| q97_7 | --- | 0.736 (fixed) | 0.180 | --- | --- | --- |
| q97_9 | --- | 0.803 (34.93) | 0.180 | --- | --- | --- |
| q97_3 | --- | -0.685 (-30.17) | 0.180 | --- | --- | --- |
| q97_4 | --- | -0.471 (-19.47) | 0.180 | --- | --- | --- |
| r88_2 | --- | --- | --- | 0.854 (fixed) | --- | 0.180 |
| r88_3 | --- | --- | --- | 0.803 (55.35) | --- | 0.180 |
| r88_6 | --- | --- | --- | 0.719 (46.60) | --- | 0.180 |
| r88_5 | --- | --- | --- | -0.789 (-48.78) | --- | 0.180 |
| r88_7 | --- | --- | --- | -0.840 (-52.94) | --- | 0.180 |
| r88_9 | --- | --- | --- | -0.730 (-41.88) | --- | 0.180 |
| r79_5 | --- | --- | --- | --- | 0.783 (fixed) | 0.180 |
| r79_7 | --- | --- | --- | --- | 0.854 (34.93) | 0.180 |
| r79_3 | --- | --- | --- | --- | -0.728 (-30.17) | 0.180 |
| r79_8 | --- | --- | --- | --- | -0.501 (-19.47) | 0.180 |

Table 4: Correlations between Content and Style Factors in Model 3 (T-Values in Parentheses).

| | THR95 | DISTR95 | STYLE95 | THR99 | DISTR99 | STYLE99 |
|---|---|---|---|---|---|---|
| THR95 | 1.000 (37.30) | | | | | |
| DISTR95 | 0.496 (15.12) | 1.000 (19.04) | | | | |
| STYLE95 | --- | --- | 1.000 (5.85) | | | |
| THR99 | **0.859** (34.79) | 0.504 (16.42) | --- | 1.000 (26.23) | | |
| DISTR99 | 0.523 (17.36) | **0.736 (18.34)** | --- | 0.592 (19.97) | 1.000 (20.68) | |
| STYLE99 | --- | --- | **0.562**(3.61) | --- | --- | 1.000 (5.36) |

# APPENDIX: MODEL SPECIFICATIONS

The measurement model for two content factors $\eta_j$ each measured with (balanced) sets of six and four indicators $y_i$ respectively, measured twice in a two wave panel, is described as follows (Bollen 1989: 151):



Where $t \in \{1995,1999\}$; $j \in \{1,2\}$ for $t = 1995$; $j \in \{3,4\}$ for $t = 1999$; $i \in \{1,...,6\}$ for $j = 1$; $i \in \{7,...,10\}$ for $j = 2$; $i \in \{11,...,16\}$ for $j = 3$; $i \in \{17,...,20\}$ for $j = 4$.

The intercepts $\tau_{i,j}$ do not play a role in the measurement model since no hypotheses are specified about the equality of intercepts or about the mean structure of the latent variables (all random variables are assumed to have zero means) (Jöreskog and Sörbom 1996: 297). Our hypothesis concerns the stability of the style factor across waves, i.e., we hypothesise a relation between latent variables, and do not assume scalar invariance of the observed indicators. For this purpose, only metric invariance is required, in which the slopes $\lambda i,j$ are subjected to equality constraints (Van de Vijver and Leung 1997). The structural equation model in this study simplifies to (Bollen 1989, p. 18):

MODEL 1. BASE MODEL WITHOUT STYLE FACTORS

The following constraints apply to the measurement model in which invariance of the slopes $\lambda_{i,j}$ (metric invariance) across time is specified. The slopes belonging to the corresponding observed indicators in 1995 and 1999 are invariant:

$$\lambda_{i,1}^{(1995)} = \lambda_{j,3}^{(1999)} \neq 0 \qquad \text{for } i \in \{1,\ldots,6\} \text{ and } j \in \{11,\ldots,16\}$$

$$\lambda_{i,2}^{(1995)} = \lambda_{j,4}^{(1999)} \neq 0 \qquad \text{for } i \in \{7,\ldots,10\} \text{ and } j \in \{17,\ldots,20\}$$

The observed indicators are not related to the latent variables that they do not measure:

$\lambda_{i,2}^{(1995)} = \lambda_{j,4}^{(1999)} = 0$ for $i \in \{1,\ldots,6\}$ and $j \in \{11,\ldots,16\}$

$\lambda_{i,1}^{(1995)} = \lambda_{j,3}^{(1999)} = 0$ for $i \in \{7,\ldots,10\}$ and $j \in \{17,\ldots,20\}$

All error variances are free:

$\mathrm{Var}(\varepsilon_i \varepsilon_j) \neq 0$ for $i \in \{1,\ldots,20\}$ and $i = j$

All error co-variances are fixed to 0, except for the co-variances between identical indicators in 1995 and 1999:

$\mathrm{Cov}(\varepsilon_i \varepsilon_j) = 0$ for $i \in \{1,...,20\}$ and $i \in \{1,...,20\}$ and $i \neq j$ and $j \neq i + 10$

$\mathrm{Cov}(\varepsilon_i \varepsilon_j) \neq 0$ for $i \in \{1,...,10\}$ and $j = i + 10$

Variances and co-variances $\psi_{i,j}$ of latent variables $\eta_j$ are different from 0:

$\psi_{i,j} \neq 0$ for $i \in \{1,4\}$ and $j \in \{1,4\}$

ADDITIONAL SPECIFICATIONS OF MODEL 2

Model 2 has the same parameter specifications of Model 1 plus additional specifications for the style factors $\eta_5$ in 1995 and $\eta_6$ in 1999.

All slopes $\lambda i,j$ of style factors to indicators are identical since it is assumed that all observed indicators are equally affected by style in 1995 and in 1999:

$\lambda_{1,5} = \lambda_{1,6}$

$\lambda_{i,5} = \lambda_{j,5}$ for $i \in \{1,\ldots,10\}$ and $j \in \{1,\ldots,10\}$ and $i \neq j$

$\lambda_{i,6} = \lambda_{j,6}$ for $i \in \{1,\ldots,10\}$ and $j \in \{1,\ldots,10\}$ and $i \neq j$

The co-variances between the style factors and the content factors are constrained to 0, the co-variance between the style factors measured in 1995 and in 1999 is also constrained to 0, but the variances of the style factors are not constrained to 0:

$\psi_{i,5}^{(1995)} = 0$ for $i \in \{1,4\}$

$\psi_{i,6}^{(1999)} = 0$ for $i \in \{1,4\}$

$\psi_{5,6} = 0$

$\psi_{i,i} \neq 0$ for $i \in \{5,6\}$ and $i = j$

ADDITIONAL SPECIFICATIONS OF MODEL 3

This model has the same specifications as in Models 1 and 2 with one exception: The covariance between the style factors measured in 1995 and 1999 is relaxed:

$\psi_{5,6} \neq 0$

FOOTNOTES

1. The surveys of the two major language groups are done by two different survey organisations located at the “Katholieke Universiteit leuven” (ISPO), and ‘L’Université Catholique de Louvain (PIOP)’
2. The measurement model with two content factors and one acquiescence factor was already identified and confirmed by testing it on two samples of the Flemish part of Belgium in 1995 and on one sample in Wallonia in 1995, and we do not repeat this procedure in the different samples. Furthermore, a test of the correlation between the acquiescence factor and a variable that counts the number of times the respondent agreed to the Likert-scale response items was already performed with the same data and variables, and the style factor was identified as an acquiescence factor. We do not repeat this procedure here, but rather refer the reader to Billiet and McClendon (2000).
3. The likelihood of at least one missing value in the two times six items about perceived ethnic threat and the two times four items of distrust in politics was separately tested in a logistic regression model with item non-response as a dependent variable and with gender, age, and education level as predictors. Two attitudinal variables with items from other parts of the questionnaire were also included as predictors in the models: an ethnic prejudice scale and a political trust scale, both with only a few missing observations. None of the attitudinal variables had any effect on the likelihood of a missing observation. However, item non-response is not independent with respect to each of the three social-background variables. Women, older people, and lower educated respondents are more likely to belong to the group with item non-response. Since acquiescence is more likely among older people and the less educated, we may assume that the acquiescence is somewhat underestimated. It was possible to retain more respondents for our analysis by means of imputation, but for our purpose, we found it much safer to analyse the actual observed responses.

4. As Model 1, which does not include style factors, is equivalent to a model with two content factors for each measurement year and style factors for each year with variances equal to zero, Model 2 can be considered as nested in Model 1. Similarly, Model 3 is nested in Model 2. For that reason, the differences in chi-square values in relation to the drop in degrees of freedom in each model provide valuable information for model choice
5. In fact, we know about one dataset in which the latter condition has been fulfilled − the 1995 ISSP data. The last item in the quasi balanced set of items about perceived ethnic threat is a five-point scale item stating that the number of immigrants should be increased or reduced. We have tested a measurement model with a style factor (acquiescence) for the whole set with samples of six Western European countries. In none of these countries was there a non-zero factor loading of the response style on the 'increase-reduce' items (Billiet, Cambré and Welkenhuysen-Gybels 2002: 64). This is the only indication we have that our style factor is acquiescence rather than a response order effect or the preference for a specific number.